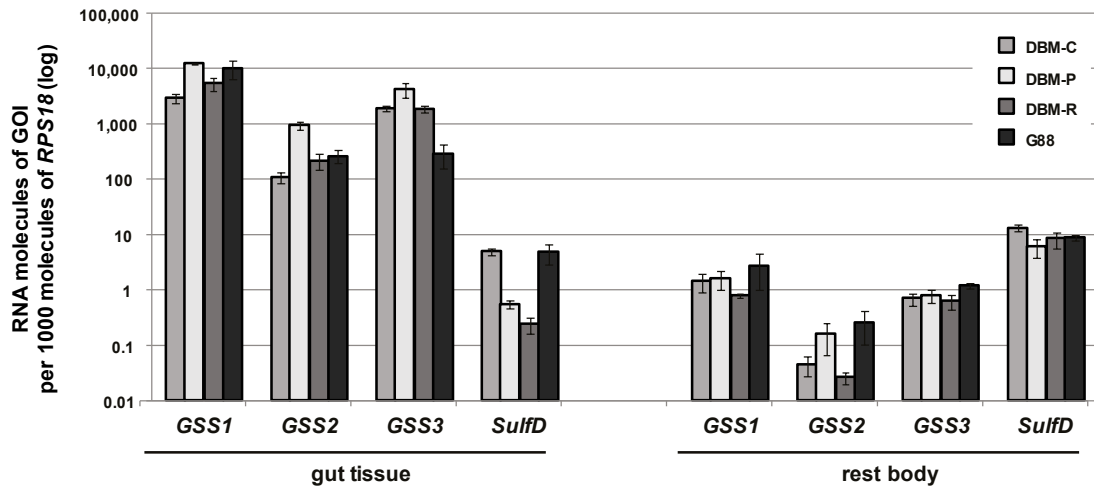


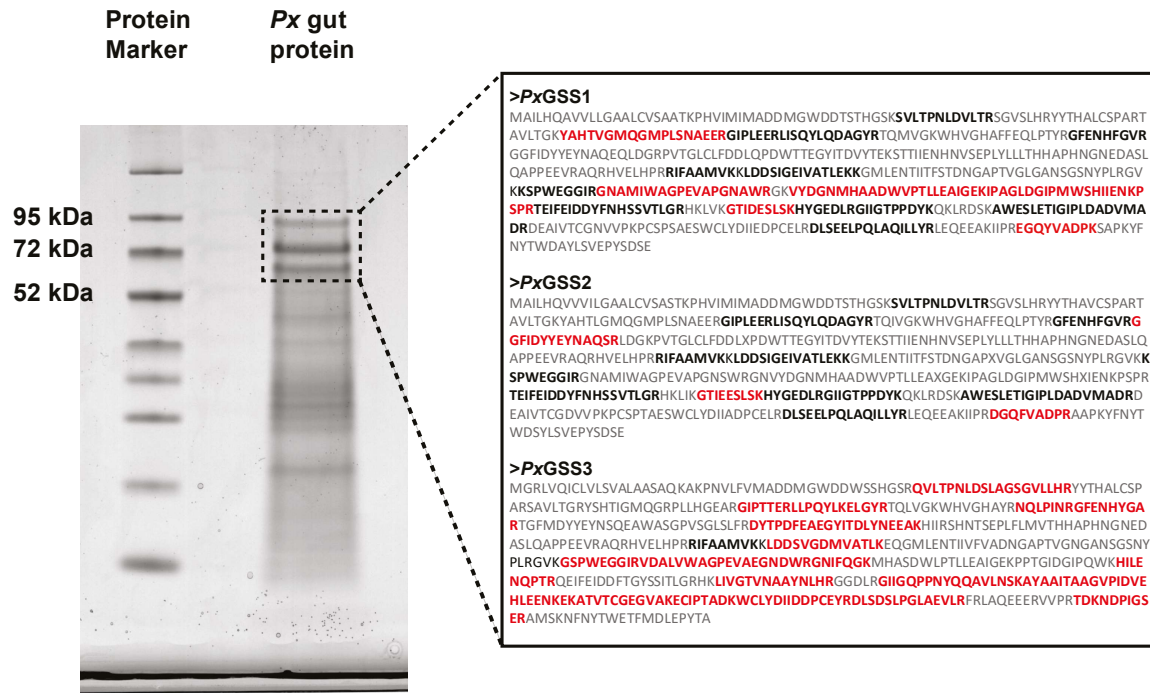
1



2

3 **Figure S1. Abundance of *PxGSS1 – 3* and *PxSulfD* transcripts in gut (left) and larval body**
 4 **(right) of four *P. xylostella* strains, DBM-C, -P, -R and G88.** Larvae were reared on *Brassica*
 5 *oleracea*, *B. napus*, *Pisum sativum*, or on standard artificial diet, respectively. Shown are results
 6 from three biological replicates of ten pooled individuals per strain (\pm SEM). Genes of interest
 7 (GOI) were compared to *RPS18* as a reference gene.

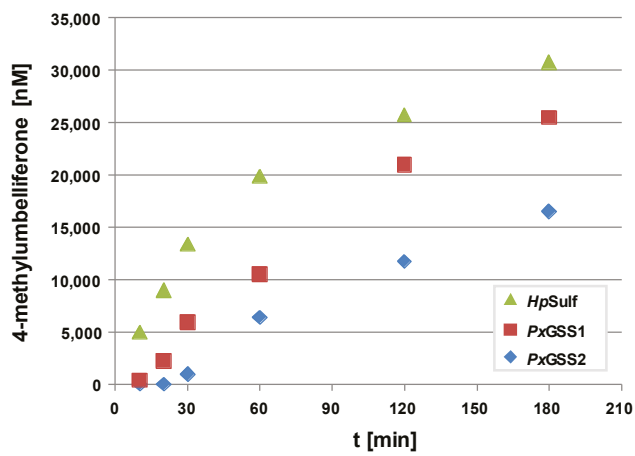
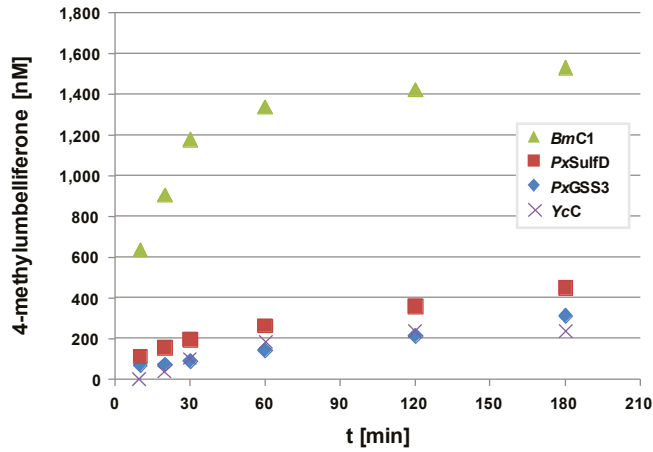
8



9

10 **Figure S2. Peptide sequencing by mass spectrometry (MS^E) of DBM-R.** Three bands (dotted
11 square region) from an SDS-PAGE protein gel (left) of *P. xylostella* strain DBM-R were
12 subjected to MS^E analysis as described in the methods. The box (right) shows GSS1 – 3 amino
13 acid sequences. Black letters correspond to peptides that are present in more than one GSS, red
14 letters correspond to unique peptides for a given GSS.

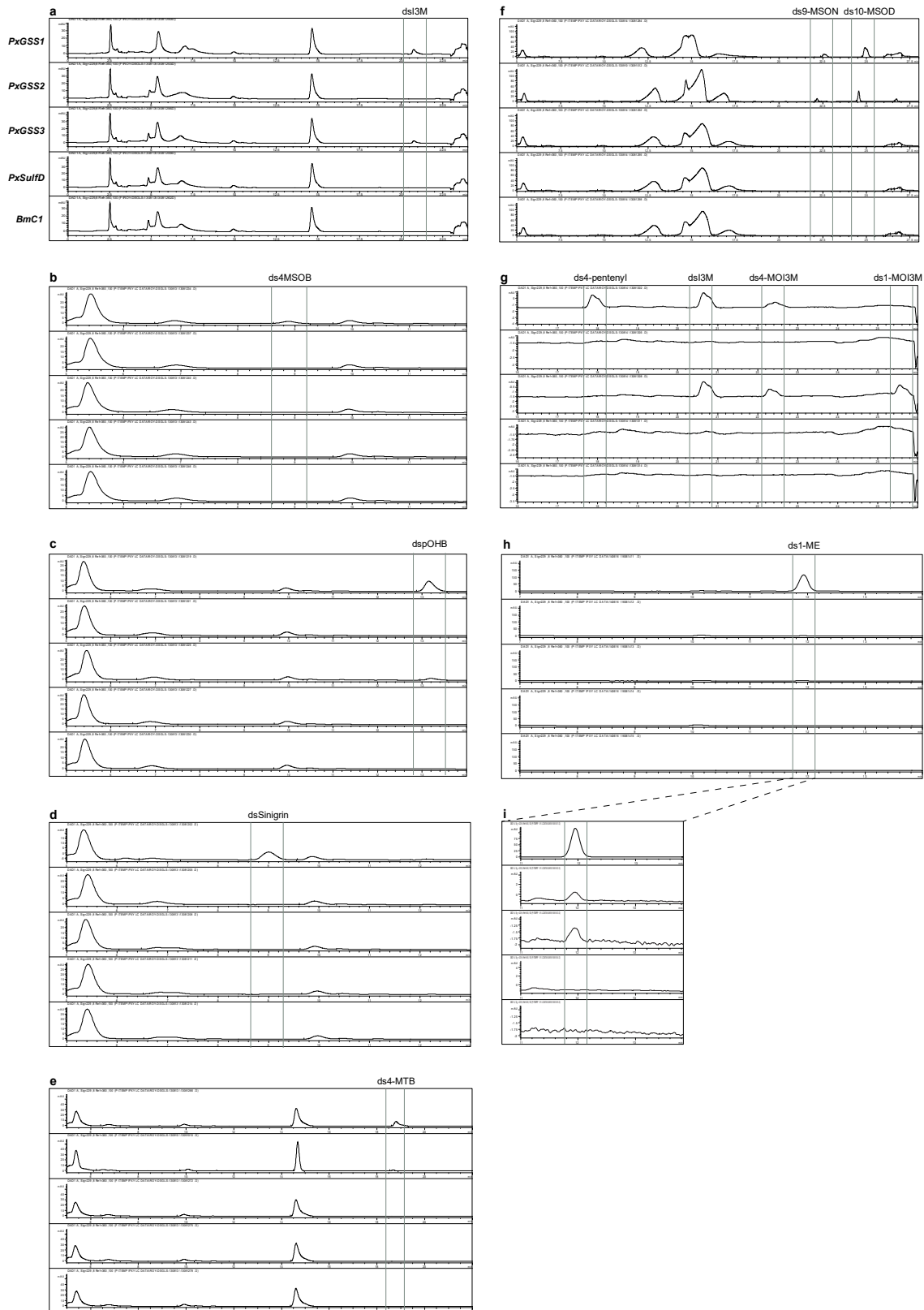
15



16

17 **Figure S3. Activity of recombinant arylsulfatases with the arylsulfatase standard**
 18 **substrate 4-methylumbelliferyl sulfate.** Heterologously expressed genes from *P. xylostella*
 19 (*PxGSS1 – 3* and *PxSulfD*), *B. mori* (*BmC1*) and *Y. cagnagella* (*YcC*) are active sulfatases.
 20 *HpSulf*, a sulfatase extracted from the burgundy snail *Helix pomatia* was used as a positive
 21 control.

22



23

24 **Figure S4. HPLC chromatograms from arylsulfatase enzyme assays.** Individual
 25 glucosinolates (a-e) or glucosinolate mixtures extracted from *Camelina sativa* (f), *Brassica*
 26 *napus* (g) or *Sisymbrium officinale* (h, i) were used for enzymatic assays with recombinant
 27 proteins (*PxGSS1*, *PxGSS2*, *PxGSS3*, *PxSulfD* and *BmC1*). Desulfo-glucosinolates were
 28 identified by HPLC at 229 nm.



30

31 **Figure S5. HPLC chromatograms from arylsulfatase enzyme assays (continued).**

32 Individual glucosinolates (a-e) or glucosinolate mixtures extracted from *Camelina sativa* (f) or

33 *Sisymbrium officinale* (g) were used for enzymatic assays with recombinant proteins (*PxGSS1*,

34 *YcC*). Desulfo-glucosinolates were identified by HPLC at 229 nm.

35

36 **Table S1. Results from branch-site tests for positive selection – Support for positive**
 37 **selection along branches.**

38

Foreground branch in Figure 5 ((((((PxGSS1a, PxGSS1b) #1, (PxGSS2a, PxGSS2b) #2) #3, PaGSS1/2 #4) #5, ((PxGSS3a, PxGSS3b) #6, PaGSS3 #7) #8) #9, YcC #10) #11, ((BmC1 #12, BmC2 #13) #14, (SfC1 #15, SfC2 #16) #17) #18), (((((PxD, PaD), YcD), (BmD, SfD)), ((PxB, YcB), (BmB, SfB))), (BmA, SfA));	$\ln L(\omega_2 = 1.0)$	$\ln L(\omega_2 \geq 1.0)$	$2\Delta l$	$p(df=1)$	adjusted ($df=1$)
#1	-21356.4	-21356.4	0.0	0.979	n.s.
#2	-21354.1	-21350.0	8.2	0.004	≤ 0.05
#3	-21355.7	-21353.5	4.3	0.038	n.s.
#4	-21353.8	-21343.7	20.2	0.000	0.000
#5	-21328.2	-21311.6	33.1	0.000	0.000
#6	-21356.4	-21354.9	3.0	0.085	n.s.
#7	-21356.4	-21356.4	0.0	1.000	n.s.
#8	-21339.6	-21322.3	34.7	0.000	0.000
#9	-21350.3	-21347.0	6.7	0.010	n.s.
#10	-21347.4	-21342.2	10.4	0.001	≤ 0.05
#11	-21356.4	-21355.8	1.2	0.269	n.s.
#12	-21354.9	-21353.6	2.5	0.113	n.s.
#13	-21356.4	-21351.8	9.1	0.003	≤ 0.05
#14	-21344.5	-21340.5	8.0	0.005	≤ 0.05
#15	-21349.1	-21348.7	0.7	0.389	n.s.
#16	-21343.9	-21340.5	6.9	0.009	n.s.
#17	-21352.5	-21348.2	8.6	0.003	≤ 0.05
#18	-21352.3	-21346.5	11.7	0.001	≤ 0.01

39

40

41

42 **Table S2. Results from branch-site tests for positive selection – Site classes.**

Foreground branch in Figure 5	site class	0	1	2a	2b
#2	proportion	0.86	0.13	0.01	0.00
	background ω	0.10	1	0.10	1
	foreground ω	0.10	1	98	98
#4	proportion	0.86	0.12	0.02	0.00
	background ω	0.10	1	0.10	1
	foreground ω	0.10	1	85	85
#5	proportion	0.76	0.10	0.12	0.02
	background ω	0.09	1	0.09	1
	foreground ω	0.09	1	56	56
#8	proportion	0.78	0.10	0.10	0.01
	background ω	0.09	1	0.09	1
	foreground ω	0.09	1	999	999
#10	proportion	0.80	0.11	0.07	0.01
	background ω	0.09	1	0.09	1
	foreground ω	0.09	1	365	365

43

44

48 **Table S4. Branch-site unrestricted statistical tests for episodic diversification (BUSTED).**

49

Branch	Model	log L	# params	AIC _c	p	Branch set	ω_1	ω_2	ω_3
(((((((PxGSS1a, PxGSS1b) #1, (PxGSS2a, PxGSS2b) #2) #3, PaGSS1/2 #4) #5, ((PxGSS3a, PxGSS3b) #6, PaGSS3 #7) #8) #9, YcC #10) #11, ((BmC1 #12, BmC2 #13) #14, (SfC1 #15, SfC2 #16) #17) #18), (((((PxD, PaD), YcD), (BmD, SfD)), ((PxB, YcB), (BmB, SfB))), (BmA, SFA));									
#1	Unconstrained	-21014	69	42168	1.000	Test	0.00 (70.12%)	0.24 (19.92%)	9.52 (9.96%)
	Constrained	-21014	68	42166		Background	0.01 (74.87%)	0.32 (19.74%)	1.84 (5.39%)
#2	Unconstrained	-21010	69	42159	0.010	Test	0.03 (74.87%)	0.07 (24.18%)	55.11 (0.95%)
	Constrained	-21015	68	42166		Background	0.01 (75.05%)	0.32 (18.80%)	1.60 (6.15%)
#3	Unconstrained	-21012	69	42163	0.117	Test	0.00 (24.77%)	0.00 (59.01%)	1.00 (16.22%)
	Constrained	-21014	68	42166		Background	0.01 (75.09%)	0.32 (18.75%)	1.60 (6.17%)
#4	Unconstrained	-21003	69	42146	0.000	Test	0.08 (13.58%)	0.09 (85.67%)	89.44 (0.74%)
	Constrained	-21014	68	42165		Background	0.01 (75.78%)	0.33 (17.96%)	1.56 (6.26%)
#5	Unconstrained	-20970	69	42079	0.000	Test	0.00 (28.48%)	0.00 (51.00%)	1.00 (20.53%)
	Constrained	-21003	68	42143		Background	0.01 (75.81%)	0.35 (19.03%)	1.84 (5.16%)
#6	Unconstrained	-21010	69	42160	0.623	Test	0.00 (15.82%)	0.00 (82.05%)	43.24 (2.13%)
	Constrained	-21011	68	42159		Background	0.01 (73.08%)	0.23 (16.99%)	1.12 (9.93%)
#7	Unconstrained	-21012	69	42164	-	Test	0.00 (4.71%)	0.00 (80.39%)	1.00 (14.91%)
						Background	0.01 (72.65%)	0.22 (17.19%)	1.11 (10.15%)
#8	Unconstrained	-20980	69	42100	0.000	Test	1.00 (56.41%)	1.00 (29.90%)	196.55 (13.69%)
	Constrained	-20996	68	42128		Background	0.01 (80.02%)	0.55 (19.66%)	122.91 (0.31%)
#9	Unconstrained	-21007	69	42153	1.000	Test	1.00 (13.69%)	1.00 (56.41%)	1.00 (29.90%)
	Constrained	-21007	68	42151		Background	0.01 (80.56%)	0.58 (19.12%)	132.77 (0.32%)
#10	Unconstrained	-21014	69	42166	0.99	Test	0.00 (85.89%)	0.01 (12.55%)	4.05 (1.56%)
	Constrained	-21014	68	42164		Background	0.01 (74.45%)	0.33 (20.07%)	1.92 (5.48%)
#11	Unconstrained	-21012	69	42164	-	Test	0.00 (13.80%)	0.00 (81.59%)	1.00 (4.60%)
						Background	0.01 (74.54%)	0.33 (20.01%)	1.92 (5.45%)
#12	Unconstrained	-20980	69	42100	0.000	Test	0.05 (100.00%)	0.71 (0.00%)	1.97 (0.00%)
	Constrained	-20996	68	42128		Background	0.01 (76.04%)	0.39 (19.74%)	2.33 (4.22%)
#13	Unconstrained	-20980	69	42100	0.000	Test	1.00 (43.20%)	1.00 (46.74%)	326.40 (10.07%)
	Constrained	-20996	68	42128		Background	0.01 (80.29%)	0.60 (19.41%)	123.19 (0.30%)
#14	Unconstrained	-21007	69	42153	1.000	Test	0.00 (64.90%)	1.00 (0.00%)	1.00 (35.10%)
	Constrained	-21007	68	42151		Background	0.01 (80.56%)	0.61 (19.15%)	133.64 (0.29%)
#15	Unconstrained	-21007	69	42153	1.000	Test	0.01 (88.32%)	1.00 (0.00%)	∞ (11.68%)
	Constrained	-21007	68	42151		Background	0.00 (65.86%)	0.24 (27.80%)	1.75 (6.34%)
#16	Unconstrained	-21014	69	42166	0.99	Test	0.00 (88.31%)	0.98 (0.00%)	1.00 (11.68%)
	Constrained	-21014	68	42164		Background	0.00 (65.76%)	0.23 (27.91%)	1.75 (6.33%)
#17	Unconstrained	-21014	69	42166	0.99	Test	0.00 (71.48%)	0.36 (27.31%)	370.60 (1.22%)
	Constrained	-21014	68	42164		Background	0.01 (74.71%)	0.30 (18.67%)	1.59 (6.62%)
#18	Unconstrained	-21014	69	42166	0.99	Test	0.00 (71.85%)	0.37 (28.15%)	1.00 (0.00%)
	Constrained	-21014	68	42164		Background	0.01 (74.71%)	0.30 (18.76%)	1.61 (6.54%)

50

51 **Table S5. Adaptive branch-site REL test (aBSREL) for episodic diversification.**

Branch (((((((PxGSS1a, PxGSS1b) #1, (PxGSS2a, PxGSS2b) #2) #3, PaGSS1/2 #4) #5, ((PxGSS3a, PxGSS3b) #6, PaGSS3 #7) #8) #9, YcC #10) #11, ((BmC1 #12, BmC2 #13) #14, (SfC1 #15, SfC2 #16) #17) #18), (((((PxD, PaD), YcD), (BmD, SfD)), ((PxB, YcB), (BmB, SfB))), (BmA, SfA));	Nucleotide GTR			Baseline MG94xREV			Full adaptive model			LRT	p-value	ω distribution over sites	
	AIC _c	log L	Parameters	AIC _c	log L	Parameters	AIC _c	log L	Parameters				
#1	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.8	-20799.5	172	0.3	0.3795	$\omega_1 = \infty$ (100%)	
#2	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.8	-20799.5	172	11.8	0.0009	$\omega_1 = 0.00$ (99%)	$\omega_2 = 55.0$ (1.1%)
#3	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.8	-20799.5	172	0.3	0.3668	$\omega_1 = 0.0595$ (99%)	$\omega_2 = 7.55$ (0.98%)
#4	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.9	-20799.6	172	16.1	0.0001	$\omega_1 = 0.00$ (98%)	$\omega_2 = 62.9$ (1.8%)
#5	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.8	-20799.6	172	10.0	0.0024	$\omega_1 = 0.601$ (88%)	$\omega_2 = 93.7$ (12%)
#6	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.9	-20799.6	172	2.0	0.1447	$\omega_1 = 0.000208$ (99%)	$\omega_2 = 7.40$ (1.2%)
#7	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.9	-20799.6	172	0.0	1	$\omega_1 = 0.0959$ (100%)	
#8	46679.8	-23286.8	53	42607.7	-21198.8	104	41949.7	-20800.0	172	5.0	0.0293	$\omega_1 = 0.0861$ (92%)	$\omega_2 = 25.8$ (7.9%)
#9	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.9	-20799.6	172	0.0	0.5	$\omega_1 = 0.00383$ (90%)	$\omega_2 = 53.1$ (10%)
#10	46679.8	-23286.8	53	42607.7	-21198.8	104	41948.9	-20799.6	172	0.0	0.4779	$\omega_1 = 0.00$ (72%)	$\omega_2 = 1.14$ (28%)

52

53 **Table S6. Detailed results of McDonald-Kreitman comparisons of *P. xylostella* GSS**
 54 **genes with reconstructed ancestral sequences.**

55

		PxGSS1 (N = 60)					PxGSS2 (N = 24)					PxGSS3 (N = 40)				
Polymorphisms		syn _{Jc} : 56	nsyn _{Jc} : 3	NI	χ^2_{Jc}	P_{Jc}	syn _{Jc} : 86	nsyn _{Jc} : 12	NI	χ^2_{Jc}	P_{Jc}	syn _{Jc} : 97	nsyn _{Jc} : 8	NI	χ^2_{Jc}	P_{Jc}
Sequence or ancestral sequence	Model for reconstruction	syn _{Jc}	nsyn _{Jc}	NI	χ^2_{Jc}	P_{Jc}	syn _{Jc}	nsyn _{Jc}	NI	χ^2_{Jc}	P_{Jc}	syn _{Jc}	nsyn _{Jc}	NI	χ^2_{Jc}	P_{Jc}
YcC	-	-	379.6	-	-	-	899.0	378.1	0.33	13.5	0.000238945	562.8	354.0	0.13	39.6	0.000000000
SulfC before separation of <i>Yponomeuta</i> and <i>Plutella</i> lineages	MO (1 ω ratio)	467.1	278.8	0.09	25.1	0.000000555	424.2	277.5	0.21	27.7	0.000000139	347.2	249.2	0.11	44.9	0.000000000
	branch-site 1 ($\omega_2 > 1$)	442.5	287.8	0.08	27.6	0.000000146	388.7	287.9	0.19	33.1	0.000000009	280.7	250.8	0.09	56.9	0.000000000
	branch-site 1 ($\omega_2 > 1$)	422.4	283.5	0.08	28.6	0.000000089	372.8	283.5	0.18	34.3	0.000000005	264.8	246.6	0.09	59.3	0.000000000
	branch-site 2 ($\omega_2 = 1$)	507.2	298.2	0.09	24.7	0.000000671	445.0	298.3	0.21	28.9	0.000000075	367.8	255.0	0.12	43.2	0.000000000
	branch-site 2 ($\omega_2 > 1$)	431.9	289.2	0.08	28.6	0.000000091	395.3	289.3	0.19	32.6	0.000000011	393.8	260.6	0.12	41.1	0.000000000
	branch-site 3 ($\omega_2 = 1$)	442.5	287.8	0.08	27.6	0.000000146	388.7	287.9	0.19	33.1	0.000000009	280.7	250.8	0.09	56.9	0.000000000
SulfC before initial gene duplication in the <i>Plutella</i> lineage (GSS1/2, GSS3)	branch-site 3 ($\omega_2 > 1$)	422.4	283.5	0.08	28.6	0.000000089	372.8	283.5	0.18	34.3	0.000000005	264.8	246.6	0.09	59.3	0.000000000
	MO (1 ω ratio)	124.3	109.7	0.06	34.8	0.000000004	124.2	110.8	0.16	36.2	0.000000002	52.2	86.9	0.05	75.7	0.000000000
	branch-site 1 ($\omega_2 = 1$)	160.2	147.7	0.06	37.6	0.000000001	158.0	148.9	0.15	40.8	0.000000000	30.9	55.0	0.05	68.0	0.000000000
	branch-site 1 ($\omega_2 > 1$)	164.2	150.1	0.06	37.4	0.000000001	162.0	151.4	0.15	40.5	0.000000000	30.9	55.0	0.05	68.0	0.000000000
	branch-site 2 ($\omega_2 = 1$)	59.9	127.2	0.03	71.2	0.000000000	65.1	129.6	0.07	77.1	0.000000000	91.1	72.4	0.10	41.0	0.000000000
	branch-site 2 ($\omega_2 > 1$)	12.3	97.0	0.01	111.2	0.000000000	18.7	100.4	0.03	111.8	0.000000000	165.9	102.8	0.13	34.0	0.000000000
GSS1/2 (before separation of <i>P. xylostella</i> and <i>P. australiana</i> lineages)	branch-site 3 ($\omega_2 = 1$)	160.2	147.7	0.06	37.6	0.000000001	158.0	148.9	0.15	40.8	0.000000000	30.9	55.0	0.05	68.0	0.000000000
	branch-site 3 ($\omega_2 > 1$)	164.2	150.1	0.06	37.4	0.000000001	162.0	151.4	0.15	40.5	0.000000000	30.9	55.0	0.05	68.0	0.000000000
	MO (1 ω ratio)	1.0	6.0	0.01	34.6	0.000000004	7.1	12.1	0.08	25.3	0.000000485					
	branch-site 1 ($\omega_2 = 1$)	3.0	5.0	0.03	22.1	0.000002603	9.2	11.1	0.12	19.3	0.000011371					
	branch-site 1 ($\omega_2 > 1$)	3.0	5.0	0.03	22.1	0.000002603	9.2	11.1	0.12	19.3	0.000011371					
	branch-site 2 ($\omega_2 = 1$)	5.1	6.0	0.04	20.2	0.000006972	11.3	12.1	0.13	18.5	0.000016708					
PxGSS1/2 (after separation of <i>P. xylostella</i> and <i>P. australiana</i> lineages)	branch-site 2 ($\omega_2 > 1$)	5.1	6.0	0.04	20.2	0.000006972	11.3	12.1	0.13	18.5	0.000016708					
	branch-site 3 ($\omega_2 = 1$)	3.0	5.0	0.03	22.1	0.000002603	9.2	11.1	0.12	19.3	0.000011371					
	branch-site 3 ($\omega_2 > 1$)	3.0	5.0	0.03	22.1	0.000002603	9.2	11.1	0.12	19.3	0.000011371					
	MO (1 ω ratio)	1.0	0.0	-	0.1	0.817922928	5.1	6.0	0.12	12.8	0.000344587					
	branch-site 1 ($\omega_2 = 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
	branch-site 1 ($\omega_2 > 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
GSS3 (before separation of <i>P. xylostella</i> and <i>P. australiana</i> lineages)	branch-site 2 ($\omega_2 = 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
	branch-site 2 ($\omega_2 > 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
	branch-site 3 ($\omega_2 = 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
	branch-site 3 ($\omega_2 > 1$)	1.0	0.0	-	-	-	5.1	6.0	0.12	12.8	0.000344587					
	MO (1 ω ratio)											35.5	5.0	0.58	0.8	0.365765569
	branch-site 1 ($\omega_2 = 1$)											30.9	7.0	0.36	3.5	0.059977742
PaGSS1/2	branch-site 1 ($\omega_2 > 1$)											30.9	7.0	0.36	3.5	0.059977742
	branch-site 2 ($\omega_2 = 1$)											22.0	6.0	0.30	4.5	0.033286162
	branch-site 2 ($\omega_2 > 1$)											22.0	6.0	0.30	4.5	0.033286162
	branch-site 3 ($\omega_2 = 1$)											30.9	7.0	0.36	3.5	0.059977742
	branch-site 3 ($\omega_2 > 1$)											30.9	7.0	0.36	3.5	0.059977742
	MO (1 ω ratio)															
branch-site models																
branch-site 1 ($\omega_2 = 1$)																
branch-site 1 ($\omega_2 > 1$)																
branch-site 2 ($\omega_2 = 1$)																
branch-site 2 ($\omega_2 > 1$)																
branch-site 3 ($\omega_2 = 1$)																
branch-site 3 ($\omega_2 > 1$)																

56

57

58 **Table S7. Primers used in this study.**

Name	Sequence (5'→3')	Usage
PXY_GSS3FpIB	GTGATGGGTCGGCTAGTGCAGATC	expression
PXY_GSS3RpIB	AGCGGTGTAAGGCTCGAGGTC	expression
PXY_SulfDpIB	GAAATGGTGTACAAATGGGCGC	expression
PXY_SulfDpIB	GAAAACATTTTGTAAAACGCATATAAACTG	expression
PXY_GSS1FpIB	GTGATGGCGATTCTGCATCAAGCTG	expression
PXY_GSS1RpIB	TTCGGAGTCGGAGTAGGGTTC	expression
PXY_GSS2FpIB	AGCATGGCAATTCTGCATCAAGTAGTGGTGATTCTG	expression
PXY_GSS2RpIB	TTCAGAGTCAGAGTAGGGTTCAACAGAG	expression
BMO_C1FpIB	GAAATGGTCGTGTTTCGAGTTTTGCTG	expression
BMO_C1RpIB	ATCAATAAATGTATCCACGTATTGTTATG	expression
YCA_CFpIB	GAAATGGTTTTGCTAGCGGCCTTTTTCG	expression
YCA_CRpIB	TTTTTTATTTATCTGATTCTTTGACTTAATGG	expression
PXY_GSS3Fq	CTGAGGTCGCTGAAGGAAAC	RT-qPCR
PXY_GSS3Rq	TGATGTCGTACAGGCACCAT	RT-qPCR
PXY_SulfDFq	TACCTCAAGGAGCTCGGCTA	RT-qPCR
PXY_SulfDRq	AATGCTCTCACCACTGACC	RT-qPCR
PXY_GSS1Fq	CATCAAGCTGTGGTGCTCCTC	RT-qPCR
PXY_GSS1Rq	CTGCCGTCAAGCTGCTCC	RT-qPCR
PXY_GSS2Fq	CTGCATCAAGTAGTGGTGATTCTG	RT-qPCR
PXY_GSS2Rq	GCTTGCCGTCAAGCCTCGAT	RT-qPCR
PXY_RPS18Fq	ATGTCGCTCGTAATACCGGAC	RT-qPCR
PXY_RPS18Rq	CAGTAGTGACGCATGCCTCG	RT-qPCR